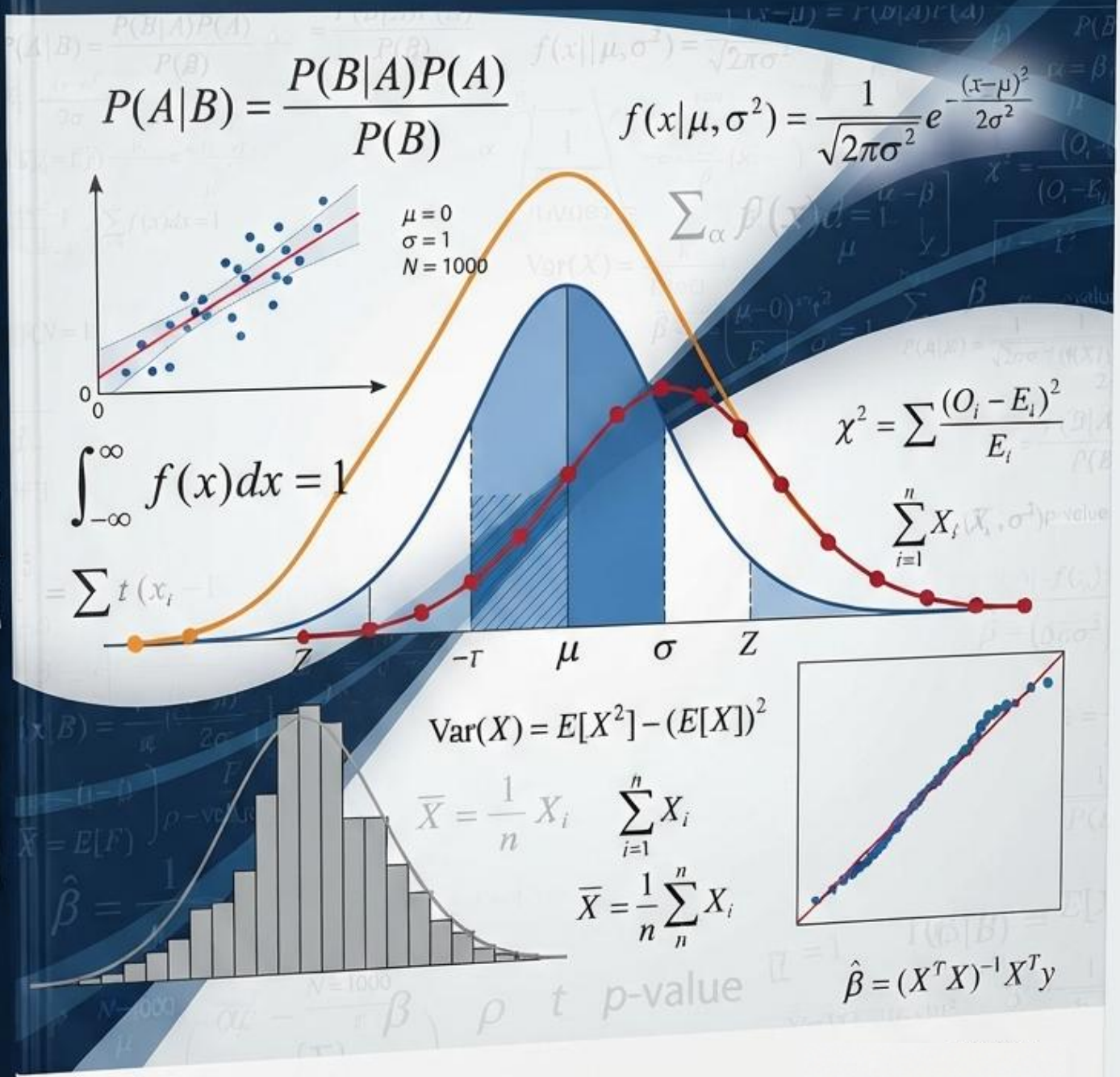
PROBABILITY AND STATISTICAL INFERENCE

NINTH EDITION



Hogg | Tanis | Zimmerman

Ninth Edition

Pearson

MAS 112

အခန်း (၁) - ဖြစ်နိုင်ခြေ (Probability)

ဤအခန်းသည် စာရင်းအင်းပညာ၏ အခြေခံအုတ်မြစ်ဖြစ်သည့် **ဖြစ်နိုင်ခြေ (Probability)** ၏ အခြေခံသဘောတရားများကို ရှင်းလင်းထားပါသည်။ ဖြစ်နိုင်ခြေဆိုသည်မှာ မသေချာမရေရာသော အခြေအနေတစ်ခုတွင် ဖြစ်ရပ်တစ်ခု ဖြစ်ပေါ်လာနိုင်သည့် အခွင့်အလမ်းကို တိုင်းတာခြင်းဖြစ်သည်။

၁.၁ ဖြစ်နိုင်ခြေ၏ ဂုဏ်သတ္တိများ (Properties of Probability)

ဤအပိုင်းတွင် ဖြစ်နိုင်ခြေ၏ အခြေခံဂုဏ်သတ္တိများကို သတ်မှတ်ပြီး ဖြစ်ရပ်များ (Events) နှင့် နမူနာနေရာ (Sample Space) ဟူသော အခြေခံအသုံးအနှုန်းများကို မိတ်ဆက်ထားပါသည်။

အဓိကအချက်များ-

- **စာရင်းအင်းပညာ (Statistics)** ဆိုသည်မှာ အချက်အလက်များကို စုဆောင်းခြင်းနှင့် ခွဲခြမ်းစိတ်ဖြာခြင်းဖြစ်သည်။ အချက်အလက်များသည် ပုံစံအမျိုးမျိုး ကွဲပြားလေ့ရှိပြီး စာရင်းအင်းပညာရှင်များသည် ဤကွဲပြားမှုကို နားလည်ကာ ပုံစံများကို ရှာဖွေကြသည်။
- **ဖြစ်နိုင်ခြေ (Probability)** သည် ဖြစ်ရပ်တစ်ခု၏ ဖြစ်ပေါ်လာနိုင်ခြေကို ၀ မှ ၁ ကြားရှိ ဂဏန်းတစ်ခုဖြင့် သတ်မှတ်သည်။
- **နမူနာနေရာ (Sample Space - S)** ဆိုသည်မှာ စမ်းသပ်မှုတစ်ခု၏ ဖြစ်နိုင်သော ရလဒ်အားလုံး၏ အစုအဝေးဖြစ်သည်။
- **ဖြစ်ရပ် (Event - A)** ဆိုသည်မှာ နမူနာနေရာ၏ အစုအဝေးတစ်ခုဖြစ်သည်။ စမ်းသပ်မှုရလဒ်သည် A တွင်ပါဝင်ပါက ဖြစ်ရပ် A ဖြစ်ပေါ်သည်ဟု ဆိုသည်။
- **ဖြစ်နိုင်ခြေသတ်မှတ်ချက် (Probability Axioms):**
    1. $P(A) \geq 0$ (ဖြစ်နိုင်ခြေသည် အနှုတ်မဖြစ်နိုင်ပါ။)
    2. $P(S) = 1$ (နမူနာနေရာတစ်ခုလုံး၏ ဖြစ်နိုင်ခြေသည် ၁ ဖြစ်သည်။)

3. အကယ်၍ $A_1, A_2, A_3, \ldots$ တို့သည် သီးခြားဖြစ်ရပ်များ (Mutually Exclusive Events) ဖြစ်ပါက၊ $P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \cdots$ ဖြစ်သည်။

အရေးကြီးသော သီအိုရီများ (Theorems):

1. **ဖြည့်စွက်ဖြစ်ရပ် (Complement):** $P(A') = 1 - P(A)$
2. **ဗလာအစု (Empty Set):** $P(\emptyset) = 0$
3. **အုပ်စုဝင်ဖြစ်ရပ် (Subset):** အကယ်၍ $A \subset B$ ဖြစ်ပါက $P(A) \leq P(B)$ ဖြစ်သည်။
4. **ပေါင်းဆုံဖြစ်ရပ် (Union):** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5. **ဖြစ်နိုင်ခြေအားလုံး၏ ပေါင်းလဒ် (General Addition Rule):** $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

ဥပမာ ၁.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
အံစာတုံးတစ်ခုကို ၆ ကြိမ်ပစ်ချပြီး အနည်းဆုံး တစ်ကြိမ်တွင် မျက်နှာနံပါတ်နှင့် အကြိမ်ရေ ကိုက်ညီမှု (Match) ရှိခြင်း ဖြစ်နိုင်ခြေကို တွက်ချက်ခြင်း။

၁.၂ ရေတွက်နည်းစနစ်များ (Methods of Enumeration)

ဤအပိုင်းတွင် ဖြစ်နိုင်ခြေများကို တွက်ချက်ရာတွင် အသုံးပြုသော ရေတွက်နည်းစနစ်များကို ရှင်းလင်းထားပါသည်။

အဓိကနည်းလမ်းများ-

1. **မြှောက်ခြင်းစည်းမျဉ်း (Multiplication Principle):** အကယ်၍ လုပ်ဆောင်ချက် E_1 တွင် n_1 နည်းလမ်းနှင့် E_2 တွင် n_2 နည်းလမ်းရှိပါက၊ လုပ်ဆောင်ချက်နှစ်ခုစလုံးကို လုပ်ဆောင်ရန် $n_1 \times n_2$ နည်းလမ်းရှိသည်။
2. **ပြောင်းလဲအစီအစဉ် (Permutation):** n မတူညီသော အရာဝတ္ထုများမှ r ခုကို ရွေးချယ်ပြီး အစီအစဉ်ကျအောင် စီရန် နည်းလမ်းအရေအတွက်-

$$^{n}P_{r} = \frac{n!}{(n-r)!}$$

3. **ပေါင်းစပ်အစီအစဉ် (Combination):** n မတူညီသော အရာဝတ္ထုများမှ r ခုကို အစီအစဉ်မပါဘဲ ရွေးချယ်ရန် နည်းလမ်းအရေအတွက်-

$$^{n}C_{r} = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

4. **ထပ်ခါထပ်ခါပြောင်းလဲအစီအစဉ် (Permutation with Repetition):** n အရာဝတ္ထုများတွင် အမျိုးအစားတူသော အရာဝတ္ထုများပါဝင်ပါက၊ စီစဉ်နိုင်သော နည်းလမ်းအရေအတွက်-

$$\frac{n!}{n_1!\,n_2!\,\ldots n_k!}$$

၁.၃ အခြေအနေအရ ဖြစ်နိုင်ခြေ (Conditional Probability)

ဖြစ်ရပ် B ဖြစ်ပွားပြီးသောအခါ ဖြစ်ရပ် A ဖြစ်ပွားနိုင်ခြေကို **အခြေအနေအရ ဖြစ်နိုင်ခြေ (Conditional Probability)** ဟုခေါ်သည်။

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)},\ P(B) > 0$$

မြှောက်ခြင်းစည်းမျဉ်း (Multiplication Rule):

$$P(A \cap B) = P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$$

ဘေးဆိုင်ရာဖြစ်ရပ်များ (Independent Events):

ဖြစ်ရပ် A နှင့် B တို့သည် အောက်ပါအခြေအနေတွင် သီးခြားဖြစ်ရပ်များ (Independent) ဖြစ်သည်။

$$P(A \cap B) = P(A) \cdot P(B)$$

ဆိုလိုသည်မှာ ဖြစ်ရပ်တစ်ခု ဖြစ်ပွားခြင်းသည် အခြားဖြစ်ရပ်၏ ဖြစ်နိုင်ခြေကို မပြောင်းလဲစေပါ။

၁.၄ သီးခြားဖြစ်ရပ်များ (Independent Events)

ဖြစ်ရပ် A နှင့် B တို့သည် သီးခြားဖြစ်ရပ်များဖြစ်ပါက၊ အောက်ပါအတွဲများလည်း သီးခြားဖြစ်ရပ်များဖြစ်သည်-

- A နှင့် B'
- A' နှင့် B
- A' နှင့် B'

အပြန်အလှန်သီးခြားဖြစ်ရပ်များ (Mutually Independent Events):
ဖြစ်ရပ် A, B, C တို့သည် အောက်ပါအခြေအနေနှစ်ခုစလုံး ပြည့်မီပါက အပြန်အလှန်သီးခြားဖြစ်ရပ်များဖြစ်သည်-

1. $P(A \cap B) = P(A)P(B),\ P(A \cap C) = P(A)P(C),\ P(B \cap C) = P(B)P(C)$
2. $P(A \cap B \cap C) = P(A)P(B)P(C)$

၁.၅ Bayes ၏ သီအိုရမ် (Bayes' Theorem)

Bayes သီအိုရမ် သည် ကြိုတင်ဖြစ်နိုင်ခြေ (Prior Probability) များကို အခြေအနေအရ ဖြစ်နိုင်ခြေများဖြင့် ပြင်ဆင်ကာ နောက်ဆက်တွဲဖြစ်နိုင်ခြေ (Posterior Probability) ကို တွက်ချက်ရန် အသုံးပြုသည်။

$$P(B_k \mid A) = \frac{P(B_k) \cdot P(A \mid B_k)}{\sum_{i=1}^{n} P(B_i) \cdot P(A \mid B_i)}$$

ဥပမာ ၁.၅-၂ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
စက်ရုံသုံးခုမှ ထုတ်လုပ်သော ပစ္စည်းတစ်ခု ချို့ယွင်းနေပါက၊ ၎င်းသည် စက်ရုံ III မှ ထုတ်လုပ်ထားခြင်း ဖြစ်နိုင်ခြေကို တွက်ချက်ခြင်း။

ဤအခန်း၏ အဓိကအချက်မှာ ဖြစ်နိုင်ခြေသည် စာရင်းအင်းပညာ၏ အခြေခံဖြစ်ပြီး မသေချာမရေရာသော အခြေအနေများတွင် ဆုံးဖြတ်ချက်များချရန် အရေးပါသော

ကိရိယာတစ်ခုဖြစ်ကြောင်း ဖော်ပြထားပါသည်။ ဖြစ်နိုင်ခြေ၏ အခြေခံဂုဏ်သတ္တိများ၊ ရေတွက်နည်းစနစ်များ၊ အခြေအနေအရ ဖြစ်နိုင်ခြေ၊ သီးခြားဖြစ်ရပ်များ၊ နှင့် Bayes သီအိုရမ်တို့သည် နောင်အခန်းများတွင် လေ့လာမည့် ကိန်းဂဏန်းဆိုင်ရာ ကောက်ချက်ချခြင်းအတွက် အခြေခံအုတ်မြစ်များဖြစ်သည်။

အခန်း (၂) - သီးခြားဖြန့်ဝေမှုများ (Discrete Distributions)

ဤအခန်းတွင် သီးခြားကျပန်းကိန်းရှင်များ (Discrete Random Variables) ၏ ဖြန့်ဝေမှုပုံစံများကို အသေးစိတ် ရှင်းလင်းထားပါသည်။ သီးခြားဖြန့်ဝေမှုဆိုသည်မှာ ကျပန်းကိန်းရှင်တစ်ခုသည် တန်ဖိုးအရေအတွက် အကန့်အသတ်ဖြင့်သာ ယူနိုင်သော ဖြန့်ဝေမှုဖြစ်သည်။

၂.၁ သီးခြားအမျိုးအစား ကျပန်းကိန်းရှင်များ (Random Variables of the Discrete Type)

သီးခြားကျပန်းကိန်းရှင် (Discrete Random Variable) ဆိုသည်မှာ ရေတွက်၍ရသော တန်ဖိုးများ (ဥပမာ - ကိန်းပြည့်များ) ကိုသာ ယူနိုင်သော ကျပန်းကိန်းရှင်ဖြစ်သည်။

ဖြစ်နိုင်ခြေဒြပ်ထုလုပ်ဆောင်ချက် (Probability Mass Function - pmf):

$$f(x) = P(X = x)$$

pmf ၏ ဂုဏ်သတ္တိများ-

1. $f(x) \geq 0$
2. $\sum_x f\,(x) = 1$
3. $P(X \in A) = \sum_{x \in A} f\,(x)$

စုစည်းဖြန့်ဝေမှုလုပ်ဆောင်ချက် (Cumulative Distribution Function - cdf):

$$F(x) = P(X \leq x) = \sum_{t \leq x} f\,(t)$$

ဥပမာ ၂.၁-၃ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

အံစာတုံး ၄ မျက်နှာပါ (tetrahedral die) ကို နှစ်ကြိမ်ပစ်ချပြီး အမြင့်ဆုံးတန်ဖိုး (maximum) ၏ ဖြန့်ဝေမှုကို ရှာဖွေခြင်း။

၂.၂ သင်္ချာဆိုင်ရာ မျှော်လင့်ချက် (Mathematical Expectation)

သခ်ၤာဆိုင်ရာ မျှော်လင့်ချက် (Mathematical Expectation) ဆိုသည်မှာ ကျပန်းကိန်းရှင်တစ်ခု၏ ပျမ်းမျှတန်ဖိုးကို တွက်ချက်ခြင်းဖြစ်သည်။

$$E(X) = \sum_x x\, f(x)$$

မျှော်လင့်ချက်၏ ဂုဏ်သတ္တိများ-

1. $E(c) = c$ (ကိန်းသေတစ်ခု၏ မျှော်လင့်ချက်သည် ထိုကိန်းသေဖြစ်သည်။)
2. $E(cX) = cE(X)$
3. $E(X + Y) = E(X) + E(Y)$

၂.၃ အထူးသခ်ၤာဆိုင်ရာ မျှော်လင့်ချက်များ (Special Mathematical Expectations)

ပျမ်းမျှ (Mean) သို့မဟုတ် ပထမလဟုတ်စဉ် (First Moment):

$$\mu = E(X)$$

ကွဲလွဲမှု (Variance) သို့မဟုတ် ဒုတိယဗဟိုလဟုတ်စဉ် (Second Central Moment):

$$\sigma^2 = E[(X - \mu)^2] = E(X^2) - \mu^2$$

စံသွေဖည်မှု (Standard Deviation):

$$\sigma = \sqrt{\sigma^2}$$

လဟုတ်စဉ်ထုတ်လုပ်သည့်လုပ်ဆောင်ချက် (Moment-Generating Function - mgf):

$$M(t) = E(e^{tX}) = \sum_x e^{tx} f(x)$$

mgf မှ လဟုတ်စဉ်များကို ရှာဖွေနည်း-

$$E(X^k) = M^{(k)}(0)$$

- $M^{(k)}(0)$ သည် $M(t)$ ၏ k-th ဆင်းသက်လက်ကို $t = 0$ တွင် တွက်ချက်ခြင်းဖြစ်သည်။

၂.၄ Binomial ဖြန့်ဝေမှု (The Binomial Distribution)

Binomial ဖြန့်ဝေမှုသည် လွတ်လပ်သော Bernoulli စမ်းသပ်မှု n ကြိမ်တွင် အောင်မြင်မှု (success) အရေအတွက်ကို ဖော်ပြသည်။

pmf-

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x},\ x = 0,1,2,\dots,n$$

ဂုဏ်သတ္တိများ-

- $E(X) = np$
- $Var(X) = np(1-p)$
- $M(t) = (1 - p + pe^t)^n$

၂.၅ Negative Binomial ဖြန့်ဝေမှု (The Negative Binomial Distribution)

Negative Binomial ဖြန့်ဝေမှုသည် r-th အောင်မြင်မှုရရှိရန် လိုအပ်သော စမ်းသပ်မှုအရေအတွက်ကို ဖော်ပြသည်။

pmf-

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r},\ x = r, r+1, r+2, \dots$$

ဂုဏ်သတ္တိများ-

- $E(X) = \frac{r}{p}$
- $Var(X) = \frac{r(1-p)}{p^2}$
- $M(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r$

၂.၆ Poisson ဖြန့်ဝေမှု (The Poisson Distribution)

Poisson ဖြန့်ဝေမှုသည် သတ်မှတ်အချိန်အတွင်း ဖြစ်ပေါ်သော ဖြစ်ရပ်အရေအတွက်ကို ဖော်ပြသည်။

pmf-

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!},\ x = 0,1,2,\dots$$

ဂုဏ်သတ္တိများ-

- $E(X) = \lambda$
- $Var(X) = \lambda$
- $M(t) = e^{\lambda(e^t-1)}$

အကျဉ်းချုပ်ဇယား (Summary Table)

ဖြန့်ဝေမှု (Distribution)	pmf	ပျမ်းမျှ (Mean)	ကွဲလွဲမှု (Variance)	mgf
Binomial	$\binom{n}{x}p^x(1-p)^{n-x}$	np	$np(1-p)$	$(1-p+pe^t)^n$
Negative Binomial	$\binom{x-1}{r-1}p^r(1-p)^{x-r}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{pe^t}{1-(1-p)e^t}\right)^r$
Poisson	$\frac{e^{-\lambda}\lambda^x}{x!}$	λ	λ	$e^{\lambda(e^t-1)}$

ဤအခန်း၏ အဓိကအချက်မှာ သီးခြားဖြန့်ဝေမှုများသည် လက်တွေ့ဘဝတွင် အသုံးအများဆုံး ဖြန့်ဝေမှုများဖြစ်ပြီး ၎င်းတို့၏ ဂုဏ်သတ္တိများကို သိရှိခြင်းသည် ကိန်းဂဏန်းဆိုင်ရာ ကောက်ချက်ချခြင်းအတွက် အရေးပါကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၃) - စဉ်ဆက်မပြတ်ဖြန့်ဝေမှုများ (Continuous Distributions)

ဤအခန်းတွင် စဉ်ဆက်မပြတ်ကျပန်းကိန်းရှင်များ (Continuous Random Variables) ၏ ဖြန့်ဝေမှုပုံစံများကို အသေးစိတ် ရှင်းလင်းထားပါသည်။ စဉ်ဆက်မပြတ်ဖြန့်ဝေမှုဆိုသည်မှာ ကျပန်းကိန်းရှင်တစ်ခုသည် ကြားကာလတစ်ခုအတွင်းရှိ မည်သည့်တန်ဖိုးကိုမဆို ယူနိုင်သော ဖြန့်ဝေမှုဖြစ်သည်။

၃.၁ စဉ်ဆက်မပြတ်အမျိုးအစား ကျပန်းကိန်းရှင်များ (Random Variables of the Continuous Type)

စဉ်ဆက်မပြတ်ကျပန်းကိန်းရှင် (Continuous Random Variable) ဆိုသည်မှာ ကြားကာလတစ်ခုအတွင်းရှိ မည်သည့်တန်ဖိုးကိုမဆို ယူနိုင်သော ကျပန်းကိန်းရှင်ဖြစ်သည်။

ဖြစ်နိုင်ခြေသိပ်သည်းဆလုပ်ဆောင်ချက် (Probability Density Function - pdf):

$$f(x) \geq 0, \int_{-\infty}^{\infty} f(x)dx = 1$$

စုစည်းဖြန့်ဝေမှုလုပ်ဆောင်ချက် (Cumulative Distribution Function - cdf):

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

မျှော်လင့်ချက် (Expectation):

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)dx$$

ကွဲလွဲမှု (Variance):

$$Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = E(X^2) - \mu^2$$

ဥပမာ ၃.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

ညီမျှဖြန့်ဝေမှု (Uniform Distribution) $U(0,1)$ ၏ pdf နှင့် cdf ကို ရှာဖွေခြင်း။

၃.၂ Exponential, Gamma, နှင့် Chi-Square ဖြန့်ဝေမှုများ (The Exponential, Gamma, and Chi-Square Distributions)

၃.၂.၁ Exponential ဖြန့်ဝေမှု (Exponential Distribution)

Exponential ဖြန့်ဝေမှုသည် Poisson ဖြစ်စဉ်တွင် ပထမဆုံးဖြစ်ရပ် ဖြစ်ပေါ်ရန် စောင့်ဆိုင်းရသည့် အချိန်ကို ဖော်ပြသည်။

pdf-

$$f(x) = \frac{1}{\theta} e^{-x/\theta},\ x \geq 0, \theta > 0$$

ဂုဏ်သတ္တိများ-

- $E(X) = \theta$
- $Var(X) = \theta^2$
- $M(t) = \frac{1}{1-\theta t},\ t < 1/\theta$

မေ့လျော့နိုင်သော ဂုဏ်သတ္တိ (Memoryless Property):

$$P(X > s + t \mid X > s) = P(X > t)$$

၃.၂.၂ Gamma ဖြန့်ဝေမှု (Gamma Distribution)

Gamma ဖြန့်ဝေမှုသည် Exponential ဖြန့်ဝေမှု၏ အထွေထွေပုံစံဖြစ်ပြီး α-th ဖြစ်ရပ် ဖြစ်ပေါ်ရန် စောင့်ဆိုင်းရသည့် အချိန်ကို ဖော်ပြသည်။

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-x/\theta},\ x \geq 0, \alpha > 0, \theta > 0$$

ဂုဏ်သတ္တိများ-

- $E(X) = \alpha\theta$
- $Var(X) = \alpha\theta^2$
- $M(t) = \frac{1}{(1-\theta t)^{\alpha}},\ t < 1/\theta$

၃.၂.၃ Chi-Square ဖြန့်ဝေမှု (Chi-Square Distribution)

Chi-Square ဖြန့်ဝေမှုသည် Gamma ဖြန့်ဝေမှု၏ အထူးကိစ္စတစ်ခုဖြစ်ပြီး $\alpha = r/2$ နှင့် $\theta = 2$ ဖြစ်သည်။

$$f(x) = \frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-x/2},\ x \geq 0$$

ဂုဏ်သတ္တိများ-

- $E(X) = r$
- $Var(X) = 2r$
- $M(t) = \frac{1}{(1-2t)^{r/2}},\ t < 1/2$

၃.၃ ပုံမှန်ဖြန့်ဝေမှု (The Normal Distribution)

ပုံမှန်ဖြန့်ဝေမှု (Normal Distribution) သည် ကိန်းဂဏန်းပညာတွင် အရေးပါဆုံး ဖြန့်ဝေမှုဖြစ်သည်။

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ -\infty < x < \infty$$

ဂုဏ်သတ္တိများ-

- $E(X) = \mu$
- $Var(X) = \sigma^2$
- $M(t) = e^{\mu t + \sigma^2 t^2/2}$

စံပုံမှန်ဖြန့်ဝေမှု (Standard Normal Distribution):

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

အရေးကြီးသော ဂုဏ်သတ္တိ-

အကယ်၍ $Z \sim N(0,1)$ ဖြစ်ပါက $Z^2 \sim \chi^2(1)$ ဖြစ်သည်။

၃.၄ နောက်ထပ်ပုံစံများ (Additional Models)

၃.၄.၁ Weibull ဖြန့်ဝေမှု (Weibull Distribution)

Weibull ဖြန့်ဝေမှုသည် ပျက်စီးမှုနှုန်း (Failure Rate) ကို လေ့လာရန် အသုံးပြုသည်။

$$f(x) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-(x/\beta)^{\alpha}},\ x \geq 0$$

၃.၄.၂ Gompertz ဖြန့်ဝေမှု (Gompertz Distribution)

Gompertz ဖြန့်ဝေမှုသည် လူသားများ၏ သေဆုံးမှုနှုန်း (Mortality Rate) ကို ဖော်ပြရန် အသုံးပြုသည်။

ပျက်စီးမှုနှုန်း (Failure Rate):

$$\lambda(x) = ae^{bx},\ a > 0, b > 0$$

အကျဉ်းချုပ်ဇယား (Summary Table)

ဖြန့်ဝေမှု (Distribution)	pdf	ပျမ်းမျှ (Mean)	ကွဲလွဲမှု (Variance)	mgf
Uniform	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$
Exponential	$\frac{1}{\theta}e^{-x/\theta}$	θ	θ^2	$\frac{1}{1-\theta t}$
Gamma	$\frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1}e^{-x/\theta}$	$\alpha\theta$	$\alpha\theta^2$	$\frac{1}{(1-\theta t)^{\alpha}}$
Chi-Square	$\frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1}e^{-x/2}$	r	$2r$	$\frac{1}{(1-2t)^{r/2}}$
Normal	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{\mu t+\sigma^2 t^2/2}$

ဤအခန်း၏ အဓိကအချက်မှာ စဉ်ဆက်မပြတ်ဖြန့်ဝေမှုများသည် လက်တွေ့ဘဝတွင် အသုံးအများဆုံး ဖြန့်ဝေမှုများဖြစ်ပြီး ၎င်းတို့၏ ဂုဏ်သတ္တိများကို သိရှိခြင်းသည် ကိန်းဂဏန်းဆိုင်ရာ ကောက်ချက်ချခြင်းအတွက် အရေးပါကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၄) - နှစ်ဘက်ဖြန့်ဝေမှုများ (Bivariate Distributions)

ဤအခန်းတွင် ကျပန်းကိန်းရှင်နှစ်ခု၏ ပေါင်းစပ်ဖြန့်ဝေမှု (Joint Distribution) အကြောင်းကို အသေးစိတ် ရှင်းလင်းထားပါသည်။ ယခင်အခန်းများတွင် ကျပန်းကိန်းရှင်တစ်ခုတည်း၏ ဖြန့်ဝေမှုကို လေ့လာခဲ့ပြီး ယခုအခန်းတွင် ကိန်းရှင်နှစ်ခုကြား ဆက်နွယ်မှုကို လေ့လာပါမည်။

၄.၁ သီးခြားအမျိုးအစား နှစ်ဘက်ဖြန့်ဝေမှုများ (Bivariate Distributions of the Discrete Type)

သီးခြားကျပန်းကိန်းရှင်နှစ်ခု X နှင့် Y ၏ ပေါင်းစပ်ဖြစ်နိုင်ခြေဖြန့်ထုလုပ်ဆောင်ချက် (Joint Probability Mass Function - pmf) ကို အောက်ပါအတိုင်း သတ်မှတ်သည်-

$$f(x,y) = P(X, = xY = y)$$

ဂုဏ်သတ္တိများ-

1. $f(x,y) \geq 0$
2. $\sum_x \sum_y f\,(x,y) = 1$
3. $P[(X,Y) \in A] = \sum_{(x,y)\in A} f\,(x,y)$

အနားသတ်ဖြစ်နိုင်ခြေဖြန့်ထုလုပ်ဆောင်ချက် (Marginal pmf):

$$f_X(x) = \sum_y f\,(x,y),\ f_Y(y) = \sum_x f\,(x,y)$$

သီးခြားဖြစ်မှု (Independence):

X နှင့် Y တို့သည် အောက်ပါအခြေအနေတွင် သီးခြားဖြစ်သည်-

$$f(x,y) = f_X(x) f_Y(y)$$

ဥပမာ ၄.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

အံစာတုံးနှစ်တုံးပစ်ချရာတွင် သေးငယ်ဆုံးနှင့် အကြီးဆုံးတန်ဖိုးများ၏ ပေါင်းစပ်ဖြန့်ဝေမှုကို ရှာဖွေခြင်း။

၄.၂ ဆက်စပ်မှုကိန်းဂဏန်း (The Correlation Coefficient)

ကိုဗေရီယံ (Covariance):

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - \mu_X\mu_Y$$

ဆက်စပ်မှုကိန်းဂဏန်း (Correlation Coefficient):

$$\rho = \frac{Cov(X,Y)}{\sigma_X\sigma_Y}$$

ဆက်စပ်မှုကိန်းဂဏန်း၏ ဂုဏ်သတ္တိများ-

- $-1 \le \rho \le 1$
- $\rho = 0$ ဆိုသည်မှာ ဆက်စပ်မှုမရှိခြင်း (Uncorrelated) ဖြစ်သည်။
- $|\rho| = 1$ ဆိုသည်မှာ ပြီးပြည့်စုံသော ဆက်စပ်မှု (Perfect Correlation) ရှိခြင်းဖြစ်သည်။

အနည်းဆုံးနှစ်ထပ်ကိန်းဆုတ်ယုတ်မျဉ်း (Least Squares Regression Line):

$$y = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)$$

၄.၃ အခြေအနေအရ ဖြန့်ဝေမှုများ (Conditional Distributions)

အခြေအနေအရ ဖြစ်နိုင်ခြေဖြန့်ထုလုပ်ဆောင်ချက် (Conditional pmf):

$$g(x \mid y) = \frac{f(x,y)}{f_Y(y)},\ f_Y(y) > 0$$
$$h(y \mid x) = \frac{f(x,y)}{f_X(x)},\ f_X(x) > 0$$

အခြေအနေအရ ပျမ်းမျှ (Conditional Mean):

$$E(Y \mid X = x) = \sum_y y\, h(y \mid x)$$

အခြေအနေအရ ကွဲလွဲမှု (Conditional Variance):

$$Var(Y \mid X = x) = \sum_y (y - E(Y \mid X = x))^2 h(y \mid x)$$

၄.၄ စဉ်ဆက်မပြတ်အမျိုးအစား နှစ်ဘက်ဖြန့်ဝေမှုများ (Bivariate Distributions of the Continuous Type)

စဉ်ဆက်မပြတ်ကျပန်းကိန်းရှင်နှစ်ခု၏ ပေါင်းစပ်ဖြစ်နိုင်ခြေသိပ်သည်းဆလုပ်ဆောင်ချက် (Joint pdf) ကို အောက်ပါအတိုင်း သတ်မှတ်သည်-

$$f(x,y) \geq 0, \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$

အနားသတ်ဖြစ်နိုင်ခြေသိပ်သည်းဆလုပ်ဆောင်ချက် (Marginal pdf):

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy,\ f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

သီးခြားဖြစ်မှု (Independence):

$$f(x,y) = f_X(x)f_Y(y)$$

၄.၅ နှစ်ဘက်ပုံမှန်ဖြန့်ဝေမှု (The Bivariate Normal Distribution)

နှစ်ဘက်ပုံမှန်ဖြန့်ဝေမှု၏ pdf မှာ-

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}$$

အရေးကြီးသော ဂုဏ်သတ္တိများ-

1. အကယ်၍ $\rho = 0$ ဖြစ်ပါက X နှင့် Y တို့သည် သီးခြားဖြစ်သည်။
2. အခြေအနေအရ ဖြန့်ဝေမှု $Y \mid X = x$ သည် ပုံမှန်ဖြန့်ဝေမှုဖြစ်သည်။

$$Y \mid X = x \sim N\left(\mu_Y, +\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\sigma_Y^2(1-\rho^2)\right)$$

အကျဉ်းချုပ်ဇယား (Summary Table)

အကြောင်းအရာ (Topic)	ဖော်မြူလာ (Formula)

ပေါင်းစပ် pmf (Joint pmf)	$f(x,y) = P(X, = xY = y)$		
အနားသတ် pmf (Marginal pmf)	$f_X(x) = \sum_y f\,(x,y),\ f_Y(y) = \sum_x f\,(x,y)$		
သီးခြားဖြစ်မှု (Independence)	$f(x,y) = f_X(x) f_Y(y)$		
ကိုဗေရီယံ (Covariance)	$Cov(X,Y) = E(XY) - \mu_X \mu_Y$		
ဆက်စပ်မှုကိန်းဂဏန်း (Correlation)	$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$		
အခြေအနေအရ ပျမ်းမျှ (Conditional Mean)	(E(Y	X=x) = \sum_{y} y h(y	x))
နှစ်ဘက်ပုံမှန် pdf (Bivariate Normal pdf)	$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}$		

ဤအခန်း၏ အဓိကအချက်မှာ ကျပန်းကိန်းရှင်နှစ်ခုကြား ဆက်နွယ်မှုကို လေ့လာရန် နှစ်ဘက်ဖြန့်ဝေမှုများသည် အရေးပါသော ကိရိယာများဖြစ်ပြီး ၎င်းတို့၏ ဂုဏ်သတ္တိများကို သိရှိခြင်းသည် ကိန်းဂဏန်းဆိုင်ရာ ကောက်ချက်ချခြင်းအတွက် အခြေခံဖြစ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၅) - ကျပန်းကိန်းရှင်များ၏ လုပ်ဆောင်ချက်များ ဖြန့်ဝေခြင်း (Distributions of Functions of Random Variables)

ဤအခန်းတွင် ကျပန်းကိန်းရှင်တစ်ခု သို့မဟုတ် တစ်ခုထက်ပိုသော ကျပန်းကိန်းရှင်များ၏ လုပ်ဆောင်ချက် (Function) တစ်ခု၏ ဖြန့်ဝေမှုကို မည်သို့ရှာဖွေရမည်ကို ရှင်းလင်းထားပါသည်။ ဤအခန်းသည် ကိန်းဂဏန်းဆိုင်ရာ ကောက်ချက်ချခြင်း (Statistical Inference) အတွက် အခြေခံကျသော သဘောတရားများကို ဖော်ပြထားပါသည်။

၅.၁ ကျပန်းကိန်းရှင်တစ်ခု၏ လုပ်ဆောင်ချက်များ (Functions of One Random Variable)

ကျပန်းကိန်းရှင် X ၏ လုပ်ဆောင်ချက် $Y = u(X)$ သည် ကျပန်းကိန်းရှင်တစ်ခုဖြစ်ပြီး ၎င်း၏ ဖြန့်ဝေမှုကို အောက်ပါနည်းလမ်းများဖြင့် ရှာဖွေနိုင်သည်-

စုစည်းဖြန့်ဝေမှုလုပ်ဆောင်ချက်နည်းလမ်း (CDF Method):

1. Y ၏ CDF ကို ရှာပါ- $F_Y(y) = P(Y \le y) = P(u(X) \le y)$
2. PDF ကို ရှာပါ- $f_Y(y) = F_Y'(y)$

ပြောင်းလဲခြင်းနည်းလမ်း (Transformation Method):

အကယ်၍ $Y = u(X)$ သည် တစ်စုံတစ်ရာသော ပြောင်းပြန်လုပ်ဆောင်ချက် (One-to-One) ရှိပါက (ဆိုလိုသည်မှာ $X = v(Y)$ ဟု ဖော်ပြနိုင်ပါက)၊

$$f_Y(y) = f_X\big(v(y)\big) \cdot | \, v'(y) \, |$$

ဥပမာ ၅.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
X သည် Gamma ဖြန့်ဝေမှုရှိပြီး $Y = e^X$ ဖြစ်ပါက Y ၏ ဖြန့်ဝေမှုကို ရှာဖွေခြင်း။

၅.၂ ကျပန်းကိန်းရှင်နှစ်ခု၏ ပြောင်းလဲခြင်းများ (Transformations of Two Random Variables)

ကျပန်းကိန်းရှင်နှစ်ခု X_1, X_2 ၏ လုပ်ဆောင်ချက်များ $Y_1 = u_1(X_1, X_2)$ နှင့် $Y_2 = u_2(X_1, X_2)$ တို့၏ ပေါင်းစပ်ဖြန့်ဝေမှုကို ရှာဖွေရန် အောက်ပါနည်းလမ်းကို အသုံးပြုသည်-

ပြောင်းလဲခြင်းနည်းလမ်း (Transformation Method):

အကယ်၍ ပြောင်းလဲခြင်းသည် တစ်စုံတစ်ရာသော ပြောင်းပြန်လုပ်ဆောင်ချက် (One-to-One) ရှိပါက (ဆိုလိုသည်မှာ $X_1 = v_1(Y_1, Y_2)$ နှင့် $X_2 = v_2(Y_1, Y_2)$ ဟု ဖော်ပြနိုင်ပါက)၊

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}\big(v_1, (y_1, y_2)v_2(y_1, y_2)\big) \cdot |J|$$

ဤတွင် J သည် **ဂျာကိုဘီယံ (Jacobian)** ဖြစ်သည်-

$$J = \begin{vmatrix} \dfrac{\partial v_1}{\partial y_1} & \dfrac{\partial v_1}{\partial y_2} \\ \dfrac{\partial v_2}{\partial y_1} & \dfrac{\partial v_2}{\partial y_2} \end{vmatrix}$$

ဥပမာ ၅.၂-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
X_1, X_2 တို့၏ ပေါင်းစပ် pdf မှ $Y_1 = X_1/X_2$ နှင့် $Y_2 = X_2$ ၏ ဖြန့်ဝေမှုကို ရှာဖွေခြင်း။

၅.၃ ကျပန်းကိန်းရှင်များစွာ (Several Random Variables)

ကျပန်းကိန်းရှင်များစွာ $X_1, X_2, \dots, X_n$ တို့သည် သီးခြားဖြစ်ပါက၊ ၎င်းတို့၏ ပေါင်းစပ် pdf သို့မဟုတ် pmf သည် တစ်ခုချင်းစီ၏ pdf သို့မဟုတ် pmf များ၏ မြှောက်လဒ်ဖြစ်သည်-

$$f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2) \dots f_n(x_n)$$

ကျပန်းနမူနာ (Random Sample):
အကယ်၍ $X_1, X_2, \dots, X_n$ တို့သည် တူညီသော ဖြန့်ဝေမှုမှ သီးခြားဖြစ်သော ကျပန်းကိန်းရှင်များဖြစ်ပါက၊ ၎င်းတို့ကို **ကျပန်းနမူနာ (Random Sample)** ဟုခေါ်သည်။

မျဉ်းကြောင်းပေါင်းစပ်မှု (Linear Combination) ၏ ပျမ်းမျှနှင့် ကွဲလွဲမှု-

$$Y = \sum_{i=1}^{n} a_i X_i$$

$$E(Y) = \sum_{i=1}^{n} a_i E(X_i)$$

$$Var(Y) = \sum_{i=1}^{n} a_i^2 \, Var(X_i) \ \left(\text{သီးခြားဖြစ်ပါက}\right)$$

၅.၄ လဟုတ်စဉ်ထုတ်လုပ်သည့်လုပ်ဆောင်ချက်နည်းလမ်း (The Moment-Generating Function Technique)

လဟုတ်စဉ်ထုတ်လုပ်သည့်လုပ်ဆောင်ချက် (MGF) ကို အသုံးပြု၍ ကျပန်းကိန်းရှင်တစ်ခု၏ ဖြန့်ဝေမှုကို ရှာဖွေနိုင်သည်။

အကယ်၍ $Y = X_1 + X_2 + \cdots + X_n$ ဖြစ်ပြီး X_i တို့သည် သီးခြားဖြစ်ပါက၊

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t) \ldots M_{X_n}(t)$$

ဥပမာ ၅.၄-၃ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

Exponential ဖြန့်ဝေမှုမှ ကျပန်းနမူနာ ၃ ခု၏ ပေါင်းလဒ်သည် Gamma ဖြန့်ဝေမှုဖြစ်ကြောင်း MGF ဖြင့် သက်သေပြခြင်း။

၅.၅ ပုံမှန်ဖြန့်ဝေမှုများနှင့် ဆက်စပ်သော ကျပန်းလုပ်ဆောင်ချက်များ (Random Functions Associated with Normal Distributions)

သီအိုရမ်-

အကယ်၍ $X_1, X_2, \ldots, X_n$ တို့သည် သီးခြားဖြစ်သော ပုံမှန်ကျပန်းကိန်းရှင်များ $X_i \sim N(\mu_i, \sigma_i^2)$ ဖြစ်ပါက၊

$$Y = \sum_{i=1}^{n} a_i \, X_i \sim N\left(\sum_{i=1}^{n} a_i \, , \mu_i \, \sum_{i=1}^{n} a_i^2 \, \sigma_i^2\right)$$

အရေးကြီးသော ရလဒ်များ-

1. **နမူနာပျမ်းမျှ (Sample Mean):**

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

2. **နမူနာကွဲလွဲမှု (Sample Variance) နှင့် Chi-Square ဖြန့်ဝေမှု-**

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

3. **Student's t ဖြန့်ဝေမှု-**

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

၅.၆ ဗဟိုကန့်သတ်သီအိုရမ် (The Central Limit Theorem - CLT)

ဗဟိုကန့်သတ်သီအိုရမ်သည် ကိန်းဂဏာန်းပညာ၏ အရေးပါဆုံးသီအိုရီများထဲမှ တစ်ခုဖြစ်သည်။

ဗဟိုကန့်သတ်သီအိုရမ် (Central Limit Theorem):
အကယ်၍ $X_1, X_2, \ldots, X_n$ တို့သည် ပျမ်းမျှ μ နှင့် ကွဲလွဲမှု σ^2 ရှိသော မည်သည့်ဖြန့်ဝေမှုမှမဆို ကျပန်းနမူနာဖြစ်ပါက၊ n ကြီးလာသည်နှင့်အမျှ

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$$

ဆိုလိုသည်မှာ n ကြီးလာလေလေ၊ နမူနာပျမ်းမျှ၏ ဖြန့်ဝေမှုသည် ပုံမှန်ဖြန့်ဝေမှုသို့ နီးကပ်လေလေဖြစ်သည်။

အသုံးချမှု-
CLT ကို အသုံးပြု၍ မည်သည့်ဖြန့်ဝေမှုမှမဆို နမူနာပျမ်းမျှအတွက် ခန့်မှန်းခြင်းနှင့် ယူဆချက်စစ်ဆေးခြင်းများ ပြုလုပ်နိုင်သည်။

၅.၇ သီးခြားဖြန့်ဝေမှုများအတွက် အနီးစပ်ဆုံးတွက်ချက်များ (Approximations for Discrete Distributions)

ဗဟိုကန့်သတ်သီအိုရမ်ကို အသုံးပြု၍ သီးခြားဖြန့်ဝေမှုများ (ဥပမာ - Binomial, Poisson) ကို ပုံမှန်ဖြန့်ဝေမှုဖြင့် အနီးစပ်ဆုံးတွက်ချက်နိုင်သည်။

Binomial ဖြန့်ဝေမှုအတွက် အနီးစပ်ဆုံးတွက်ချက်ခြင်း-

အကယ်၍ $X \sim Binomial(n, p)$ ဖြစ်ပြီး n ကြီးပါက၊

$$X \approx N\big(n, pnp(1-p)\big)$$

မျဉ်းဆက်ပြင်ခြင်း (Continuity Correction):

သီးခြားဖြန့်ဝေမှုကို စဉ်ဆက်မပြတ်ဖြန့်ဝေမှုဖြင့် အနီးစပ်ဆုံးတွက်ချက်ရာတွင် 0.5 ကို အသုံးပြုခြင်း။

၅.၈ Chebyshev ၏ မညီမျှမှုနှင့် ဖြစ်နိုင်ခြေဆိုင်ရာ ဆုံဆည်းမှု (Chebyshev's Inequality and Convergence in Probability)

Chebyshev ၏ မညီမျှမှု (Chebyshev's Inequality):

မည်သည့်ဖြန့်ဝေမှုအတွက်မဆို၊

$$P(|\ X - \mu\ | \geq k\sigma) \leq \frac{1}{k^2}$$

ဖြစ်နိုင်ခြေဆိုင်ရာ ဆုံဆည်းမှု (Convergence in Probability):

နမူနာအရွယ်အစား n ကြီးလာသည်နှင့်အမျှ၊ နမူနာပျမ်းမျှ $\bar{X}$ သည် လူဦးရေပျမ်းမျှ μ သို့ ဖြစ်နိုင်ခြေဆိုင်ရာ ဆုံဆည်းသည်။

$$\bar{X} \xrightarrow{p} \mu$$

ဤအခန်း၏ အဓိကအချက်မှာ ကျပန်းကိန်းရှင်များ၏ လုပ်ဆောင်ချက်များ၏ ဖြန့်ဝေမှုကို ရှာဖွေနည်းများ၊ ဗဟိုကန့်သတ်သီအိုရမ်၊ နှင့် အနီးစပ်ဆုံးတွက်ချက်နည်းများသည် ကိန်းဂဏန်းဆိုင်ရာ ကောက်ချက်ချခြင်းအတွက် အခြေခံကျသော ကိရိယာများဖြစ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၆) - အမှတ်ခန့်မှန်းခြေ (Point Estimation)

ဤအခန်းတွင် ကိန်းဂဏန်းပညာ၏ အခြေခံအကျဆုံး ခန့်မှန်းခြေနည်းလမ်းများဖြစ်သည့် **အမှတ်ခန့်မှန်းခြေ (Point Estimation)** အကြောင်းကို အသေးစိတ် ရှင်းလင်းထားပါသည်။ ခန့်မှန်းခြေဆိုသည်မှာ နမူနာမှ အချက်အလက်များကို အသုံးပြု၍ လူဦးရေ၏ မသိသော ပါရာမီတာ (Parameter) များကို ခန့်မှန်းခြင်းဖြစ်သည်။

၆.၁ ဖော်ပြသော ကိန်းဂဏန်းများ (Descriptive Statistics)

ဒေတာများကို ဖော်ပြရန် အောက်ပါ ကိန်းဂဏန်းများကို အသုံးပြုသည်-

နမူနာပျမ်းမျှ (Sample Mean):

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

နမူနာကွဲလွဲမှု (Sample Variance):

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

နမူနာစံသွေဖည်မှု (Sample Standard Deviation):

$$s = \sqrt{s^2}$$

ပင်ကိုယ်ဖြန့်ဝေမှု (Empirical Distribution):
ဒေတာတစ်ခုစီအတွက် အလေးချိန် $1/n$ ပေးခြင်းဖြင့် ဖန်တီးထားသော ဖြန့်ဝေမှု။

၆.၂ စူးစမ်းရှာဖွေသော ဒေတာခွဲခြမ်းစိတ်ဖြာခြင်း (Exploratory Data Analysis - EDA)

EDA သည် ဒေတာ၏ ပုံစံများကို ရှာဖွေရန် ဂရုပ်ဖစ်နှင့် ကိန်းဂဏန်းနည်းလမ်းများကို အသုံးပြုသည်။

အဓိကနည်းလမ်းများ-

- **ပင်စည်နှင့်အရွက်ပုံ (Stem-and-Leaf Display):** ဒေတာကို မူရင်းတန်ဖိုးများဖြင့် ဖော်ပြသည်။

- **သေတ္တာပုံ (Box-and-Whisker Plot):** ဒေတာ၏ ဖြန့်ကျဲမှုနှင့် အစွန်းထွက်တန်ဖိုးများ (Outliers) ကို ဖော်ပြသည်။
- **ရာခိုင်နှုန်းများ (Percentiles):** ဒေတာ၏ ရာခိုင်နှုန်းအလိုက် တန်ဖိုးများ။

၆.၃ အစီအစဉ်စာရင်းအင်းကိန်းများ (Order Statistics)

အစီအစဉ်စာရင်းအင်းကိန်းများသည် နမူနာကို အငယ်ဆုံးမှ အကြီးဆုံးသို့ စီထားသော တန်ဖိုးများဖြစ်သည်။

$$Y_1 < Y_2 < \cdots < Y_n$$

r-th အစီအစဉ်စာရင်းအင်းကိန်း၏ pdf-

$$g_r(y) = \frac{n!}{(r-1)!\,(n-r)!}[F(y)]^{r-1}[1-F(y)]^{n-r}f(y)$$

၆.၄ များများဖြစ်နိုင်ဆုံး ခန့်မှန်းခြေနည်းလမ်း (Maximum Likelihood Estimation - MLE)

MLE သည် နမူနာမှ လေ့လာတွေ့ရှိရသော ဒေတာများကို ဖြစ်ပေါ်စေနိုင်ခြေ အများဆုံးဖြစ်စေသော ပါရာမီတာ၏ တန်ဖိုးကို ရှာဖွေသည်။

များများဖြစ်နိုင်ဆုံးလုပ်ဆောင်ချက် (Likelihood Function):

$$L(\theta) = \prod_{i=1}^{n} f\,(x_i;\theta)$$

MLE ရှာဖွေနည်း-

1. $L(\theta)$ ကို ရေးပါ။
2. $\ln L\,(\theta)$ ကို ရေးပါ။
3. $\frac{d}{d\theta}\ln L\,(\theta) = 0$ ကို ဖြေရှင်းပါ။

ဥပမာ ၆.၄-၃ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

ပုံမှန်ဖြန့်ဝေမှု $N(\mu, \sigma^2)$ အတွက် MLE ရှာဖွေခြင်း။

၆.၅ ရိုးရှင်းသော ဆုတ်ယုတ်မှုပြဿနာ (A Simple Regression Problem)

ရိုးရှင်းသော ဆုတ်ယုတ်မှုပုံစံ (Simple Linear Regression Model):

$$Y_i = \alpha + \beta x_i + \varepsilon_i,\ \varepsilon_i \sim N(0, \sigma^2)$$

အနည်းဆုံးနှစ်ထပ်ကိန်း ခန့်မှန်းချက်များ (Least Squares Estimates):

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$
$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$$

MLE ခန့်မှန်းချက်များ-

- $\hat{\alpha}, \hat{\beta}$ သည် အနည်းဆုံးနှစ်ထပ်ကိန်း ခန့်မှန်းချက်များနှင့် တူညီသည်။
- $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2$

၆.၆ MLE များ၏ အနီးစပ်ဆုံးဖြန့်ဝေမှုများ (Asymptotic Distributions of Maximum Likelihood Estimators)

ပုံမှန်အခြေအနေများအောက်တွင် MLE သည် အနီးစပ်ဆုံး ပုံမှန်ဖြန့်ဝေမှုရှိသည်-

$$\hat{\theta} \approx N\left(\theta, \frac{1}{nI(\theta)}\right)$$

ဤတွင် $I(\theta)$ သည် **Fisher အချက်အလက် (Fisher Information)** ဖြစ်သည်-

$$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln f\,(X;\theta)\right]$$

၆.၇ လုံလောက်သော စာရင်းအင်းကိန်းများ (Sufficient Statistics)

လုံလောက်သော စာရင်းအင်းကိန်း (Sufficient Statistic) ဆိုသည်မှာ နမူနာတွင်ပါဝင်သော ပါရာမီတာအကြောင်း အချက်အလက်အားလုံးကို ပိုင်ဆိုင်ထားသော စာရင်းအင်းကိန်းဖြစ်သည်။

Fisher-Neyman ပိုင်းခြားခြင်းသီအိုရမ် (Factorization Theorem):
$T(X)$ သည် θ အတွက် လုံလောက်သော စာရင်းအင်းကိန်းဖြစ်ပါက၊

$$f(x_1, x_2, \dots, x_n, \theta) = g(T; (x)\theta) \cdot h(x_1, x_2, \dots, x_n)$$

ဥပမာ ၆.၇-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
Poisson ဖြန့်ဝေမှုအတွက် $\bar{X}$ သည် လုံလောက်သော စာရင်းအင်းကိန်းဖြစ်ကြောင်း သက်သေပြခြင်း။

၆.၈ Bayesian ခန့်မှန်းခြေ (Bayesian Estimation)

Bayesian ခန့်မှန်းခြေတွင် ပါရာမီတာ θ ကို ကျပန်းကိန်းရှင်အဖြစ် သတ်မှတ်ပြီး ၎င်း၏ ကြိုတင်ဖြန့်ဝေမှု (Prior Distribution) $\pi(\theta)$ ကို သတ်မှတ်သည်။

နောက်ဆက်တွဲ ဖြန့်ဝေမှု (Posterior Distribution):

$$\pi(\theta \mid x) = \frac{f(x \mid \theta)\pi(\theta)}{\int f(x \mid \theta)\pi(\theta)d\theta}$$

Bayesian ခန့်မှန်းချက် (Bayesian Estimator):
နောက်ဆက်တွဲ ဖြန့်ဝေမှု၏ ပျမ်းမျှ (Mean of Posterior Distribution) ကို ခန့်မှန်းချက်အဖြစ် အသုံးပြုသည်။

အကျဉ်းချုပ်ဇယားအကြောင်းအရာဖော်မြူလာနမူနာပျမ်းမျှနမူနာကွဲလွဲမှုဆုတ်ယုတ်မှုကိန်းအချက်အလက်လုံလောက်သောစာရင်းအင်းကိန်းခန့်မှန်းချက်ဤအခန်း၏အဓိကအချက်မှာအမှတ်ခန့်မှန်းခြေသည်လူဦးရေ၏ပါရာမီတာများကိုနမူနာမှခန့်မှန်းရန်အခြေခံနည်းလမ်းဖြစ်ပြီးဆုတ်ယုတ်မှုခွဲခြမ်းစိတ်ဖြာခြင်းလုံလောက်သောစာရင်းအင်းကိန်းများနှင့်ခန့်မှန်းခြေတို့သည်အရေးပါသောကိရိယာများဖြစ်ကြောင်းဖော်ပြထားပါသည်။

အခန်း (၇) - ကြားကာလခန့်မှန်းခြေ (Interval Estimation)

ဤအခန်းတွင် ကိန်းဂဏန်းပညာ၏ အရေးပါသော အပိုင်းတစ်ခုဖြစ်သည့် **ကြားကာလခန့်မှန်းခြေ (Interval Estimation)** အကြောင်းကို အသေးစိတ် ရှင်းလင်းထားပါသည်။ ယုံကြည်မှုကြားကာလ (Confidence Interval) တည်ဆောက်ပုံ၊ အဓိပ္ပာယ်ဖွင့်ဆိုပုံနှင့် အသုံးချမှုများကို ဖော်ပြထားပါသည်။

၇.၁ ပျမ်းမျှအတွက် ယုံကြည်မှုကြားကာလ (Confidence Intervals for Means)

ယုံကြည်မှုကြားကာလ (Confidence Interval) ဆိုသည်မှာ လူဦးရေ၏ ပါရာမီတာတစ်ခု (ဥပမာ - ပျမ်းမျှ) ပါဝင်နိုင်သည့် တန်ဖိုးအပိုင်းအခြားကို သတ်မှတ်ယုံကြည်မှုအဆင့် (Confidence Level) ဖြင့် ခန့်မှန်းခြင်းဖြစ်သည်။

၇.၁.၁ ကွဲလွဲမှု (σ^2) ကို သိသော အခါ (When Variance is Known)

လူဦးရေ၏ ကွဲလွဲမှု σ^2 ကို သိပါက၊ ပျမ်းမျှ μ အတွက် $100(1-\alpha)\%$ ယုံကြည်မှုကြားကာလမှာ-

$$\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

ဥပမာ ၇.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
မီးသီး ၂၇ လုံး၏ ပျမ်းမျှသက်တမ်းအတွက် ၉၅% ယုံကြည်မှုကြားကာလ တွက်ချက်ခြင်း။

၇.၁.၂ ကွဲလွဲမှု (σ^2) ကို မသိသော အခါ (When Variance is Unknown)

လူဦးရေ၏ ကွဲလွဲမှု σ^2 ကို မသိပါက၊ နမူနာကွဲလွဲမှု S^2 ကို အသုံးပြုပြီး t-ဖြန့်ဝေမှုဖြင့် ယုံကြည်မှုကြားကာလကို တွက်ချက်သည်။

$$\bar{X} \pm t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}$$

၇.၂ ပျမ်းမျှနှစ်ခုကြား ကွာခြားချက်အတွက် ယုံကြည်မှုကြားကာလ (Confidence Intervals for the Difference of Two Means)

၇.၂.၁ ကွဲလွဲမှုများကို သိသော အခါ (When Variances are Known)

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

၇.၂.၂ ကွဲလွဲမှုများကို မသိဘဲ နမူနာအရွယ်အစား ကြီးသော အခါ (When Variances are Unknown but Sample Sizes are Large)

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

၇.၂.၃ ကွဲလွဲမှုများ တူညီသည်ဟု ယူဆနိုင်သော အခါ (When Variances are Assumed Equal)

$$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2, n_1+n_2-2} \cdot S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

- $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ (ပေါင်းစပ်ကွဲလွဲမှု)

၇.၂.၄ ကွဲလွဲမှုများ မတူညီသော အခါ (When Variances are Not Equal)

Welch ၏ အနီးစပ်ဆုံး t-စစ်ဆေးမှုကို အသုံးပြုသည်။

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

ဒီဂရီအက်ဆစ် (Degrees of Freedom) ကို အောက်ပါအတိုင်း တွက်ချက်သည်-

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}}$$

၇.၂.၅ မှီခိုနေသော နမူနာများအတွက် (For Dependent Samples)

နမူနာနှစ်ခုသည် မှီခိုနေပါက (ဥပမာ - တူညီသော လူတစ်ဦးကို အကြိမ်နှစ်ကြိမ် တိုင်းတာခြင်း)၊ ကွာခြားချက် $D_i = X_{1i} - X_{2i}$ ကို အသုံးပြုသည်။

\bar{D} \pm t_{\alpha/2, n-1} \frac{S_D}{\sqrt{n}} ``` --- ## ၇.၃ အချိုးအစားအတွက် ယုံကြည်မှုကြားကာလ (Confidence Intervals for Proportions) ### ၇.၃.၁ အချိုးအစားတစ်ခုအတွက် (For a Single Proportion) \[\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}

၇.၃.၂ အချိုးအစားနှစ်ခုကြား ကွာခြားချက်အတွက် (For the Difference of Two Proportions)

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

၇.၄ နမူနာအရွယ်အစား သတ်မှတ်ခြင်း (Sample Size Determination)

၇.၄.၁ ပျမ်းမျှအတွက် (For a Mean)

$$n = \left(\frac{Z_{\alpha/2}\sigma}{E}\right)^2$$

- E = အမှားအနားသတ် (Margin of Error)

၇.၄.၂ အချိုးအစားအတွက် (For a Proportion)

$$n = \frac{Z_{\alpha/2}^2 p(1-p)}{E^2}$$

အကယ်၍ p ၏ တန်ဖိုးကို မသိပါက $p = 0.5$ ကို အသုံးပြုနိုင်သည် (၎င်းသည် ကွဲလွဲမှုကို အမြင့်ဆုံးဖြစ်စေသောကြောင့် ရှေးရိုးဆန်သော နည်းလမ်းဖြစ်သည်)။

၇.၅ ရာခိုင်နှုန်းများအတွက် ဖြန့်ဝေမှုမပါသော ယုံကြည်မှုကြားကာလများ (Distribution-Free Confidence Intervals for Percentiles)

ဤနည်းလမ်းသည် လူဦးရေ၏ ဖြန့်ဝေမှုပုံစံကို မသိဘဲ ရာခိုင်နှုန်းများ (Percentiles) အတွက် ယုံကြည်မှုကြားကာလကို တည်ဆောက်ရန် အသုံးပြုသည်။

အဆင့်များ-

၁။ နမူနာကို စီရန်လိုအပ်သည်။
၂။ သင့်လျော်သော အစီအစဉ်စာရင်းအင်းကိန်းများ (Order Statistics) ကို ရွေးချယ်သည်။
၃။ နှစ်ထပ်ကိန်းဖြန့်ဝေမှု (Binomial Distribution) ကို အသုံးပြု၍ ယုံကြည်မှုအဆင့်ကို တွက်ချက်သည်။

၇.၆ ပိုမိုသော ဆုတ်ယုတ်မှုခွဲခြမ်းစိတ်ဖြာခြင်း (More Regression)

ဤအပိုင်းတွင် ဆုတ်ယုတ်မှုခွဲခြမ်းစိတ်ဖြာခြင်းအတွက် ယုံကြည်မှုကြားကာလများကို ရှင်းလင်းထားပါသည်။

ဆုတ်ယုတ်မှုကိန်းများအတွက် ယုံကြည်မှုကြားကာလ-

$$\hat{\beta}_j \pm t_{\alpha/2,n-k-1} \cdot SE(\hat{\beta}_j)$$

ခန့်မှန်းချက်ကြားကာလ (Prediction Interval):

$$\hat{Y}_0 \pm t_{\alpha/2,n-k-1} \cdot \sqrt{\hat{\sigma}^2(1 + x_0^T(X^TX)^{-1}x_0)}$$

၇.၇ ပြန်လည်နမူနာယူခြင်းနည်းလမ်းများ (Resampling Methods)

ပြန်လည်နမူနာယူခြင်းနည်းလမ်းများ (Resampling Methods) သည် စာရင်းအင်းတစ်ခု၏ ဖြန့်ဝေမှုကို ခန့်မှန်းရန် ကွန်ပျူတာကို အသုံးပြုသည်။

ဘွတ်စ်ထရပ်နည်းလမ်း (Bootstrap Method):

၁။ မူရင်းနမူနာမှ ပြန်လည်နမူနာယူခြင်း (Sampling with Replacement)။
၂။ စာရင်းအင်းကို တွက်ချက်ခြင်း။

၃။ အဆင့် ၁ နှင့် ၂ ကို အကြိမ်များစွာ ပြုလုပ်ခြင်း။
၄။ ခန့်မှန်းချက်၏ ဖြန့်ဝေမှုကို ရရှိခြင်း။

ဤအခန်း၏ အဓိကအချက်မှာ ကြားကာလခန့်မှန်းခြေသည် လူဦးရေ၏ ပါရာမီတာများကို တန်ဖိုးအပိုင်းအခြားဖြင့် ခန့်မှန်းရန် အရေးပါသော ကိရိယာဖြစ်ပြီး ယုံကြည်မှုအဆင့်၊ နမူနာအရွယ်အစားနှင့် အမှားအနားသတ်တို့ကို ထည့်သွင်းစဉ်းစားရန် လိုအပ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၈) - ယူဆချက်ဆိုင်ရာ စာရင်းအင်းစစ်ဆေးမှုများ (Tests of Statistical Hypotheses)

ဤအခန်းတွင် ယူဆချက်စစ်ဆေးခြင်း (Hypothesis Testing) ၏ အခြေခံသဘောတရားများ၊ စစ်ဆေးမှုအဆင့်များ၊ အမှိုးအစားအမှိုးမှိုး၊ စစ်ဆေးမှု၏ စွမ်းအား (Power)၊ နှင့် အကောင်းဆုံးပယ်ချသည့်နယ်ပယ်များ (Best Critical Regions) အကြောင်းကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၈.၁ ပျမ်းမျှတစ်ခုအကြောင်း စစ်ဆေးခြင်း (Tests About One Mean)

ယူဆချက်စစ်ဆေးခြင်းတွင် အောက်ပါအဆင့်များ ပါဝင်သည်-

၁။ သုညယူဆချက် (Null Hypothesis) H_0 နှင့် အစားထိုးယူဆချက် (Alternative Hypothesis) H_1 ကို သတ်မှတ်ခြင်း။

၂။ အရေးပါမှုအဆင့် (Level of Significance) α ကို ရွေးချယ်ခြင်း။

၃။ သင့်လျော်သော စာရင်းအင်းစစ်ဆေးမှု (Test Statistic) ကို တွက်ချက်ခြင်း။

၄။ ပယ်ချသည့်နယ်ပယ် (Critical Region) ကို သတ်မှတ်ခြင်း။

၅။ ဆုံးဖြတ်ချက်ချခြင်းနှင့် အနက်အဓိပ္ပာယ်ဖွင့်ဆိုခြင်း။

၈.၁.၁ ကွဲလွဲမှု (σ^2) ကို သိသော အခါ (When Variance is Known)

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic):

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

ဆုံးဖြတ်ချက်စည်းမျဉ်း (Decision Rule):

ယူဆချက်	ပယ်ချသည့်နယ်ပယ်		
$H_1: \mu > \mu_0$	$Z > Z_\alpha$		
$H_1: \mu < \mu_0$	$Z < -Z_\alpha$		
$H_1: \mu \neq \mu_0$	(	Z	> Z_{\alpha/2})

ဥပမာ ၈.၁-၁ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

သံမဏိဘားတစ်ခု၏ ဆန့်နိုင်အားကို စစ်ဆေးခြင်း။

၈.၁.၂ ကွဲလွဲမှု (σ^2) ကို မသိသော အခါ (When Variance is Unknown)

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic):

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

ဆုံးဖြတ်ချက်စည်းမျဉ်း (Decision Rule):

ယူဆချက်	ပယ်ချသည့်နယ်ပယ်		
$H_1: \mu > \mu_0$	$t > t_{\alpha,n-1}$		
$H_1: \mu < \mu_0$	$t < -t_{\alpha,n-1}$		
$H_1: \mu \neq \mu_0$	(	T	> t_{\alpha/2, n-1})

၈.၂ ပျမ်းမျှနှစ်ခုကြား ကွာခြားချက်ကို စစ်ဆေးခြင်း (Tests of the Equality of Two Means)

၈.၂.၁ ကွဲလွဲမှုများကို သိသော အခါ (When Variances are Known)

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

၈.၂.၂ ကွဲလွဲမှုများကို မသိဘဲ နမူနာအရွယ်အစား ကြီးသော အခါ (When Variances are Unknown but Sample Sizes are Large)

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

၈.၂.၃ ကွဲလွဲမှုများ တူညီသည်ဟု ယူဆနိုင်သော အခါ (When Variances are Assumed Equal)

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

၈.၃ အချိုးအစားများအကြောင်း စစ်ဆေးခြင်း (Tests About Proportions)

၈.၃.၁ အချိုးအစားတစ်ခုအတွက် (For a Single Proportion)

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

၈.၃.၂ အချိုးအစားနှစ်ခုကြား ကွာခြားချက်အတွက် (For the Difference of Two Proportions)

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- $\hat{p} = \frac{x_1+x_2}{n_1+n_2}$

၈.၄ Wilcoxon စစ်ဆေးမှုများ (The Wilcoxon Tests)

Wilcoxon စစ်ဆေးမှုများသည် ဖြန့်ဝေမှုပုံစံကို မသိဘဲ အသုံးပြုနိုင်သော စစ်ဆေးမှုများ (Non-parametric Tests) ဖြစ်သည်။

၈.၄.၁ Wilcoxon Signed-Rank စစ်ဆေးမှု (Wilcoxon Signed-Rank Test)

ဤစစ်ဆေးမှုသည် ပျမ်းမျှ (median) တစ်ခုကို စစ်ဆေးရန် အသုံးပြုသည်။

အဆင့်များ-

၁။ ကွာခြားချက်များ $D_i = X_i - \mu_0$ ကို တွက်ချက်ပါ။

၂။ $|D_i|$ ကို အစဉ်လိုက်စီပါ။

၃။ အဆင့်သတ်မှတ်ပါ (Rank)။

၄။ အပြုသဘောဆောင်သော အဆင့်များ၏ ပေါင်းလဒ် W ကို တွက်ချက်ပါ။

၈.၄.၂ Wilcoxon Rank-Sum စစ်ဆေးမှု (Wilcoxon Rank-Sum Test)

ဤစစ်ဆေးမှုသည် လွတ်လပ်သော နမူနာနှစ်ခု၏ ပျမ်းမျှများကို နှိုင်းယှဉ်ရန် အသုံးပြုသည်။

အဆင့်များ-

၁။ နမူနာနှစ်ခုကို ပေါင်းစပ်ပြီး အစဉ်လိုက်စီပါ။

၂။ အဆင့်သတ်မှတ်ပါ (Rank)။

၃။ နမူနာတစ်ခု၏ အဆင့်ပေါင်းလဒ် W ကို တွက်ချက်ပါ။

၈.၅ စစ်ဆေးမှု၏ စွမ်းအား (Power of a Statistical Test)

စစ်ဆေးမှု၏ စွမ်းအား (Power of a Test) ဆိုသည်မှာ သုညယူဆချက် H_0 မှားနေချိန်တွင် ၎င်းကို ပယ်ချနိုင်သည့် ဖြစ်နိုင်ခြေဖြစ်သည်။

$$\text{Power} = 1 - \beta = P(\text{Reject } H_0 \mid H_0 \text{ is false})$$

စွမ်းအားလုပ်ဆောင်ချက် (Power Function):

$$K(\theta) = P(\text{Reject } H_0 \mid \theta)$$

၈.၆ အကောင်းဆုံးပယ်ချသည့်နယ်ပယ်များ (Best Critical Regions)

Neyman-Pearson Lemma-

$H_0: \theta = \theta_0$ နှင့် $H_1: \theta = \theta_1$ အတွက် အကောင်းဆုံးပယ်ချသည့်နယ်ပယ်မှာ-

$$\frac{L(\theta_0)}{L(\theta_1)} \leq k$$

- $L(\theta)$ = များများဖြစ်နိုင်ဆုံးလုပ်ဆောင်ချက် (Likelihood Function)

- k = သတ်မှတ်ထားသော ကိန်းသေ

၈.၇ များများဖြစ်နိုင်ဆုံးအချိုးစစ်ဆေးမှု (Likelihood Ratio Tests)

များများဖြစ်နိုင်ဆုံးအချိုး (Likelihood Ratio):

$$\lambda = \frac{\max_{\theta \in \omega} L(\theta)}{\max_{\theta \in \Theta} L(\theta)}$$

- ω = H_0 အောက်ရှိ ပါရာမီတာနယ်ပယ်
- Θ = စုစုပေါင်းပါရာမီတာနယ်ပယ်

λ သည် သေးငယ်ပါက H_0 ကို ပယ်ချသည်။

အကျဉ်းချုပ်ဇယား (Summary Table)

စစ်ဆေးမှု (Test)	စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic)	အသုံးပြုသည့်ဖြန့်ဝေမှု (Distribution)
One-Sample Z-Test	$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$N(0,1)$
One-Sample t-Test	$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$t(n-1)$
Two-Sample Z-Test	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$N(0,1)$
Two-Sample t-Test	$t = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$t(n_1 + n_2 - 2)$
One-Proportion Z-Test	$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$N(0,1)$

Two-Proportion Z-Test	$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	$N(0,1)$

ဤအခန်း၏ အဓိကအချက်မှာ ယူဆချက်စစ်ဆေးခြင်းသည် နမူနာမှ အချက်အလက်များကို အသုံးပြု၍ လူဦးရေတစ်ခုလုံးအကြောင်း ဆုံးဖြတ်ချက်ချရန်အတွက် အရေးပါသော ကိရိယာဖြစ်ပြီး မှန်ကန်သော ယူဆချက်များ၊ သင့်လျော်သော စစ်ဆေးမှုများနှင့် အနက်အဓိပ္ပာယ်ဖွင့်ဆိုချက်များ လိုအပ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၉) - နောက်ထပ်စစ်ဆေးမှုများ (More Tests)

ဤအခန်းတွင် Chi-Square စစ်ဆေးမှုများ၊ ကွဲလွဲမှုခွဲခြမ်းစိတ်ဖြာခြင်း (Analysis of Variance - ANOVA)၊ နှင့် စာရင်းအင်းအရည်အသွေးထိန်းချုပ်ခြင်း (Statistical Quality Control) တို့ကို အသေးစိတ် ရှင်းလင်းထားပါသည်။

၉.၁ Chi-Square Goodness-of-Fit စစ်ဆေးမှု (Chi-Square Goodness-of-Fit Test)

ဤစစ်ဆေးမှုသည် လေ့လာထားသော အကြိမ်ရေ (Observed Frequency) နှင့် မျှော်လင့်ထားသော အကြိမ်ရေ (Expected Frequency) တို့ကို နှိုင်းယှဉ်ခြင်းဖြင့် ဒေတာတစ်ခုသည် သတ်မှတ်ထားသော ဖြန့်ဝေမှုတစ်ခုနှင့် ကိုက်ညီမှုရှိမရှိကို စစ်ဆေးသည်။

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic):

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

- O_i = အုပ်စု i တွင် လေ့လာရရှိသော အကြိမ်ရေ
- E_i = အုပ်စု i တွင် မျှော်လင့်ထားသော အကြိမ်ရေ
- k = အုပ်စုအရေအတွက်

ဒီဂရီအက်ဆစ် (Degrees of Freedom): $df = k - 1$ (ပါရာမီတာများကို ခန့်မှန်းထားပါက $df = k - 1 - m$)

ဥပမာ ၉.၁-၂ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
ဒင်္ဂါးပစ်စမ်းသပ်မှုတစ်ခု၏ ရလဒ်များသည် Binomial ဖြန့်ဝေမှုနှင့် ကိုက်ညီမှုရှိမရှိကို စစ်ဆေးခြင်း။

၉.၂ ဆက်စပ်ဇယားများ (Contingency Tables)

ဆက်စပ်ဇယားများသည် အမျိုးအစားနှစ်ခု သို့မဟုတ် နှစ်ခုထက်ပိုသော အမျိုးအစားများကြား အမှီအခိုကင်းမှုကို စစ်ဆေးရန် အသုံးပြုသည်။

စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic):

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

- O_{ij} = အတန်း i နှင့် ကော်လံ j တွင် လေ့လာရရှိသော အကြိမ်ရေ
- $E_{ij} = \frac{(Row\ Total_i)(Column\ Total_j)}{Grand\ Total}$
- r = အတန်းအရေအတွက်
- c = ကော်လံအရေအတွက်

ဒီဂရီအက်ဆစ် (Degrees of Freedom): $df = (r-1)(c-1)$

ဥပမာ ၉.၂-၃ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):
ကျောင်းသားများ၏ ကျားမနှင့် ဌာနရွေးချယ်မှုကြား အမှီအခိုကင်းမှုကို စစ်ဆေးခြင်း။

၉.၃ တစ်ချက်ကွဲလွဲမှုခွဲခြမ်းစိတ်ဖြာခြင်း (One-Factor Analysis of Variance - ANOVA)

တစ်ချက် ANOVA သည် အုပ်စုသုံးခု သို့မဟုတ် သုံးခုထက်ပိုသော အုပ်စုများ၏ ပျမ်းမျှများကို နှိုင်းယှဉ်ရန် အသုံးပြုသည်။

စုစုပေါင်းနှစ်ထပ်ကိန်းပေါင်းလဒ် (Total Sum of Squares - SST):

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{..})^2$$

အုပ်စုကြားနှစ်ထပ်ကိန်းပေါင်းလဒ် (Between-Groups Sum of Squares - SSB):

$$SSB = \sum_{i=1}^{k} n_i \, (\bar{X}_{i.} - \bar{X}_{..})^2$$

အုပ်စုအတွင်းနှစ်ထပ်ကိန်းပေါင်းလဒ် (Within-Groups Sum of Squares - SSW):

$$SSW = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2$$

ANOVA ဇယား (ANOVA Table):

အရင်းအမြစ် (Source)	SS	df	MS	F
အုပ်စုကြား (Between)	SSB	k-1	MSB = SSB/(k-1)	F = MSB/MSW
အုပ်စုအတွင်း (Within)	SSW	N-k	MSW = SSW/(N-k)	
စုစုပေါင်း (Total)	SST	N-1		

ဥပမာ ၉.၃-၂ (စာအုပ်တွင် ဖော်ပြထားသော ဥပမာ):

မော်တော်ကားပြတင်းပေါက်များအတွက် ဆွဲအားစစ်ဆေးမှုရလဒ်များကို နှိုင်းယှဉ်ခြင်း။

၉.၄ နှစ်ချက်ကွဲလွဲမှုခွဲခြမ်းစိတ်ဖြာခြင်း (Two-Way Analysis of Variance - ANOVA)

နှစ်ချက် ANOVA သည် အကြောင်းရင်းနှစ်ခု (Factors) ၏ သက်ရောက်မှုကို တစ်ပြိုင်နက်တည်း လေ့လာရန် အသုံးပြုသည်။

ပုံစံ (Model):

$$X_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

- α_i = အကြောင်းရင်း A ၏ သက်ရောက်မှု
- β_j = အကြောင်းရင်း B ၏ သက်ရောက်မှု

- $(\alpha\beta)_{ij}$ = အပြန်အလှန်သက်ရောက်မှု (Interaction Effect)

ANOVA ဇယား (ANOVA Table):

အရင်းအမြစ် (Source)	SS	df	MS	F
အကြောင်းရင်း A	SSA	a-1	MSA = SSA/(a-1)	F_A = MSA/MSE
အကြောင်းရင်း B	SSB	b-1	MSB = SSB/(b-1)	F_B = MSB/MSE
အပြန်အလှန် (Interaction)	SSAB	(a-1)(b-1)	MSAB = SSAB/((a-1)(b-1))	F_AB = MSAB/MSE
အမှား (Error)	SSE	ab(n-1)	MSE = SSE/(ab(n-1))	
စုစုပေါင်း (Total)	SST	abn-1		

၉.၅ စာရင်းအင်းအရည်အသွေးထိန်းချုပ်ခြင်း (Statistical Quality Control)

စာရင်းအင်းအရည်အသွေးထိန်းချုပ်ခြင်းသည် ထုတ်ကုန်တစ်ခု၏ အရည်အသွေးကို ထိန်းသိမ်းရန် စာရင်းအင်းနည်းလမ်းများကို အသုံးပြုသည်။

၉.၅.၁ ထိန်းချုပ်ဇယားများ (Control Charts)

$\bar{X}$ ဇယား (Chart for the Mean):

- ဗဟိုမျဉ်း (Center Line) = $\bar{\bar{X}}$
- အထက်ထိန်းချုပ်ကန့်သတ်ချက် (Upper Control Limit - UCL) = $\bar{\bar{X}} + A_2\bar{R}$
- အောက်ထိန်းချုပ်ကန့်သတ်ချက် (Lower Control Limit - LCL) = $\bar{\bar{X}} - A_2\bar{R}$

R ဇယား (Chart for the Range):

- ဗဟိုမျဉ်း (Center Line) = $\bar{R}$
- UCL = $D_4\bar{R}$

- LCL = $D_3\bar{R}$

p ဇယား (Chart for Proportion Defective):

- ဗဟိုမျဉ်း (Center Line) = $\bar{p}$
- UCL = $\bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$
- LCL = $\bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$

c ဇယား (Chart for Number of Defects):

- ဗဟိုမျဉ်း (Center Line) = $\bar{c}$
- UCL = $\bar{c} + 3\sqrt{\bar{c}}$
- LCL = $\bar{c} - 3\sqrt{\bar{c}}$

အကျဉ်းချုပ်ဇယား (Summary Table)

စစ်ဆေးမှု (Test)	စစ်ဆေးမှုစာရင်းအင်းကိန်း (Test Statistic)	အသုံးပြုသည့်ဖြန့်ဝေမှု (Distribution)
Chi-Square Goodness-of-Fit	$\chi^2 = \sum \frac{(O-E)^2}{E}$	$\chi^2(k-1)$
Chi-Square Test of Independence	$\chi^2 = \sum\sum \frac{(O-E)^2}{E}$	$\chi^2\big((r-1)(c-1)\big)$
One-Way ANOVA F-Test	$F = \frac{MSB}{MSW}$	$F(k, -1N-k)$
Two-Way ANOVA F-Test	$F = \frac{MSA}{MSE}$ သို့ $F = \frac{MSB}{MSE}$	$F(d, f_1 df_2)$

ဤအခန်း၏ အဓိကအချက်မှာ Chi-Square စစ်ဆေးမှုများ၊ ANOVA၊ နှင့် စာရင်းအင်းအရည်အသွေးထိန်းချုပ်ခြင်းတို့သည် ဒေတာခွဲခြမ်းစိတ်ဖြာခြင်းနှင့် အရည်အသွေးမြှင့်တင်ခြင်းအတွက် အရေးပါသော ကိရိယာများဖြစ်ကြောင်း ဖော်ပြထားပါသည်။

အခန်း (၁၀) - Epilogue (နိဂုံးချုပ်)

ဤအခန်းသည် စာအုပ်၏ နိဂုံးချုပ်အခန်းဖြစ်ပြီး စာရင်းအင်းပညာ၏ အရေးပါမှု၊ လက်တွေ့ဘဝတွင် အသုံးချပုံ၊ နှင့် စာရင်းအင်းပညာရှင်များ၏ အခန်းကဏ္ဍကို ခြုံငုံသုံးသပ်ထားပါသည်။

စာရင်းအင်းပညာ၏ အရေးပါမှု (The Importance of Statistics)

စာရင်းအင်းပညာသည် အချက်အလက်များကို စုဆောင်းခြင်းနှင့် ခွဲခြမ်းစိတ်ဖြာခြင်းဆိုင်ရာ ပညာရပ်ဖြစ်သည်။ နည်းပညာတိုးတက်လာမှုကြောင့် အချက်အလက်များ ပိုမိုများပြားလာသဖြင့် စာရင်းအင်းပညာရှင်များ၏ လိုအပ်ချက်သည် တစ်နေ့တခြား ကြီးထွားလာသည်။

စာရင်းအင်းပညာ၏ အဓိကအချက်များ-

1. **အချက်အလက်များကို အဓိပ္ပာယ်ဖွင့်ဆိုခြင်း:** အချက်အလက်များသည် အချက်အလက်များသာဖြစ်သည်။ ၎င်းတို့ကို အဓိပ္ပာယ်ဖွင့်ဆိုနိုင်မှသာ အချက်အလက်များ ဖြစ်လာသည်။
2. **ကွဲပြားမှုကို နားလည်ခြင်း:** ဒေတာအားလုံးတွင် ကွဲပြားမှု (Variation) ရှိသည်။ စာရင်းအင်းပညာရှင်၏ အလုပ်မှာ ဤကွဲပြားမှုကို နားလည်ပြီး ပုံစံများကို ရှာဖွေခြင်းဖြစ်သည်။
3. **ဆုံးဖြတ်ချက်ချခြင်း:** စာရင်းအင်းပညာသည် အချက်အလက်များကို အခြေခံ၍ ဆုံးဖြတ်ချက်များချရာတွင် အထောက်အကူပြုသည်။

စာရင်းအင်းပညာ၏ အသုံးချမှုများ (Applications of Statistics)

စာရင်းအင်းပညာကို အောက်ပါနယ်ပယ်များတွင် အသုံးပြုသည်-

- စာရင်းကိုင် (Accounting)
- အာမခံသင်္ချာ (Actuarial Science)

- ဇီဝဗေဒ (Biology)
- စီးပွားရေး (Economics)
- ပညာရေး (Education)
- ပတ်ဝန်းကျင်သိပ္ပံ (Environmental Science)
- ဘဏ္ဍာရေး (Finance)
- မျိုးရိုးဗေဒ (Genetics)
- ကျန်းမာရေး (Healthcare)
- ဈေးကွက်ရှာဖွေရေး (Marketing)
- ဆေးပညာ (Medicine)
- စိတ်ပညာ (Psychology)
- အားကစား (Sports)

စာရင်းအင်းဆိုင်ရာ တွေးခေါ်မှု (Statistical Thinking)

စာရင်းအင်းဆိုင်ရာ တွေးခေါ်မှုတွင် အောက်ပါအချက်များ ပါဝင်သည်-

1. **ကွဲပြားမှုသည် နေရာတိုင်းတွင် ရှိသည် (Variation is everywhere):** စာရင်းအင်းပညာရှင်များသည် ဤကွဲပြားမှုကို နားလည်ရန် ကြိုးစားကြသည်။
2. **ကွဲပြားမှုကို လျှော့ချခြင်းသည် အောင်မြင်မှု၏ သော့ချက်ဖြစ်သည် (Understanding and reducing variation is key to success):** ထုတ်ကုန်များ ပိုမိုတသမတ်တည်းဖြစ်စေရန် ကွဲပြားမှုကို လျှော့ချရန် လိုအပ်သည်။
3. **အလုပ်အားလုံးသည် အပြန်အလှန်ဆက်နွယ်နေသော လုပ်ငန်းစဉ်များ၏ စနစ်တစ်ခုတွင် ဖြစ်ပေါ်သည် (All work occurs in a system of interconnected**

processes): ဌာနအသီးသီးသည် အကောင်းဆုံးထုတ်ကုန်များနှင့် ဝန်ဆောင်မှုများကို ဖန်တီးရန် အတူတကွ လုပ်ဆောင်ရမည်။

စာရင်းအင်းပညာရှင်များ၏ အခန်းကဏ္ဍ (The Role of Statisticians)

စာရင်းအင်းပညာရှင်များသည် အောက်ပါတို့ကို ဆောင်ရွက်သည်-

1. **ပြဿနာကို သတ်မှတ်ခြင်း:** လေ့လာရမည့် ပြဿနာကို သတ်မှတ်ခြင်း။
2. **တိုင်းတာမှုများ ဖန်တီးခြင်း:** ပြဿနာကို နားလည်ရန် သင့်လျော်သော တိုင်းတာမှုများကို ဖန်တီးခြင်း။
3. **အချက်အလက်များ စုဆောင်းခြင်း:** စစ်တမ်းများ၊ စမ်းသပ်မှုများမှ အချက်အလက်များကို စုဆောင်းခြင်း။
4. **အချက်အလက်များကို အကျဉ်းချုပ်ဖော်ပြခြင်း:** ဖော်ပြသော ကိန်းဂဏန်းများနှင့် ဂရပ်ဖစ်များကို အသုံးပြု၍ အချက်အလက်များကို အကျဉ်းချုပ်ဖော်ပြခြင်း။
5. **ခွဲခြမ်းစိတ်ဖြာခြင်းနှင့် ကောက်ချက်ချခြင်း:** အချက်အလက်များကို ခွဲခြမ်းစိတ်ဖြာပြီး ကောက်ချက်များချခြင်း။
6. **အစီရင်ခံစာတင်ပြခြင်းနှင့် အကြံပြုချက်များ:** ရလဒ်များကို အစီရင်ခံစာတင်ပြပြီး အကြံပြုချက်များ ပေးခြင်း။

နိဂုံး (Conclusion)

စာရင်းအင်းပညာသည် စိတ်လှုပ်ရှားဖွယ်ရာ ပညာရပ်တစ်ခုဖြစ်ပြီး လူမှုရေး၊ ကျန်းမာရေး၊ နှင့် ရူပဗေဒဆိုင်ရာ နယ်ပယ်များတွင် အသုံးချမှုများစွာ ရှိသည်။ စာရင်းအင်းဆိုင်ရာ တွေးခေါ်မှုသည် ကွဲပြားမှုကို နားလည်ရန် အရေးပါပြီး နေ့စဉ်ဘဝတွင် ကြီးမားသော သက်ရောက်မှုရှိနိုင်သည်။

ဤအခန်း၏ အဓိကအချက်မှာ စာရင်းအင်းပညာသည် အချက်အလက်များကို အဓိပ္ပာယ်ဖွင့်ဆိုရန်၊ ကွဲပြားမှုကို နားလည်ရန်၊ နှင့် ဆုံးဖြတ်ချက်များချရန် အရေးပါသော ပညာရပ်ဖြစ်ပြီး နယ်ပယ်အသီးသီးတွင် အသုံးချနိုင်ကြောင်း ဖော်ပြထားပါသည်။